IN THE CLAIMS:

1. (Currently amended) A method for extracting visemes from an audio speech signal, comprising:

receiving successive frames of digitized analog speech information obtained from the audio speech signal at a fixed rate;

filtering each of the successive frames of digitized analog speech information to synchronously generate time domain frame classification vectors at the fixed rate, wherein each of the time domain frame classification vectors is derived from one of the successive frames of digitized analog speech information, comprising

converting each of the successive frames of digitized analog speech information to a spectral domain vector using N multi-taper discrete prolate spheroid sequence basis (MTDPSSB) functions that are factors of a Fredholm integral of the first kind, wherein N is a positive integers; and

converting each spectral domain vector to one of the time domain frame classification vectors using Inverse Discrete Cosine Transformation, wherein N is; and

synchronously generating a sequence of a set of visemes wherein each set of visemes in the sequence is derived from a corresponding one of the time domain frame classification vectors, wherein each set of visemes is synchronously generated with a latency less than 10 milliseconds with reference to a successive frame of digitized analog speech information with which the set of visemes corresponds.

2. (Currently amended) The method for extracting visemes from an audio speech signal according to claim 1, wherein in the step of synchronously generating ~~analyzing~~, each set of visemes is generated with a latency less than 100 milliseconds with reference to a successive frame of digitized analog speech information with which the set of visemes corresponds.

Claim 3 is canceled

4. (Previously presented) The method for extracting visemes from an audio speech signal according to claim 1, wherein each set of visemes includes a subset of visemes identifiers and a one to one corresponding subset of confidence numbers.

5. (Previously presented) The method for extracting visemes from an audio speech signal according to claim 1, wherein the set of visemes consists of an identity of one most likely viseme.

Claim 6 is canceled

7. (Previously presented) The method for extracting visemes from an audio speech signal according to claim 6, wherein the conversion of each of the successive frames of digitized analog speech information to a spectral domain vector comprises:

multiplying a successive frame of digitized analog speech information by one of the N MTDPSSB functions to generate N product sets of the successive frame of digitized analog speech information;

performing a fast Fourier transform (FFT) of each of the N product sets to generate N FFT sets of the successive frame of digitized analog speech information; and

combining together the N FFT sets of the successive frame of digitized analog speech information to generate a summed FFT set of the successive frame of digitized analog speech information.

8. (Previously presented) The method for extracting visemes from an audio speech signal according to claim 7, wherein the conversion of each of the successive frames of digitized analog speech information to a spectral domain vector further comprises scaling the summed FFT set of the successive frame of digitized analog speech information.

9. (Currently amended) The method for extracting visemes from an audio speech signal according to claim 1, wherein the step of synchronously generating ~~analyzing~~ comprises a spatial classification.

10. (Currently amended) The method for extracting visemes from an audio speech signal according to claim 1, wherein the step of synchronously generating ~~analyzing~~ is performed by one of a neural network and a fuzzy logic function.

11. (Previously presented) The method for extracting visemes from an audio speech signal according to claim 10, wherein the neural network is a feed-forward memory-less perceptron type neural classifier.

12. (Currently amended) An apparatus for extracting visemes from an audio speech signal, comprising:

at least one processor; and

at least one memory that stores programmed instructions that control the at least one processor to

receive successive frames of digitized analog speech information from the audio speech signal at a fixed rate,

filter each of the successive frames of digitized analog speech information to synchronously generate time domain frame classification vectors at the fixed rate, wherein each of the time domain frame classification vectors is derived from one of the successive frames of digitized analog speech information, comprising

converting each of the successive frames of digitized analog speech information to a spectral domain vector using N multi-taper discrete prolate spheroid sequence basis (MTDPSSB) functions that are factors of a Fredholm integral of the first kind, wherein N is a positive integers; and

converting each spectral domain vector to one of the time domain frame classification vectors using Inverse Discrete Cosine Transformation, wherein N is; and

synchronously generate a sequence of a set of visemes wherein each set of visemes in the sequence is derived from a corresponding one of the time domain frame classification vectors, wherein each set of visemes is synchronously generated with a latency less than 10 milliseconds with reference to a successive frame of digitized analog speech information with which the set of visemes corresponds.

13. (Currently amended) A speech receiving device, comprising:

at least one processor;

at least one memory that stores programmed instructions that control the at least one processor to

receive successive frames of digitized analog speech information from an audio speech signal at a fixed rate,

filter each of the successive frames of digitized analog speech information to synchronously generate time domain frame classification vectors at the fixed rate, wherein each of the time domain frame classification vectors is derived from one of the successive frames of digitized analog speech information, comprising

converting each of the successive frames of digitized analog speech information to a spectral domain vector using N multi-taper discrete prolate spheroid sequence basis (MTDPSSB) functions that are factors of a Fredholm integral of the first kind, wherein N is a positive integers; and

converting each spectral domain vector to one of the time domain frame classification vectors using Inverse Discrete Cosine Transformation, wherein N is; and

synchronously generate a sequence of a set of visemes wherein each set of visemes in the sequence is derived from a corresponding one of the time domain frame classification vectors, wherein each set of visemes is synchronously generated with a latency

less than 10 milliseconds with reference to a successive frame of digitized analog speech information with which the set of visemes corresponds; and

a display that displays an avatar that is formed using the set of visemes.

14. (Currently Amended) An apparatus for extracting visemes from an audio speech signal, comprising:

means for receiving successive frames of digitized analog speech information from the audio speech signal at a fixed rate,

means for filtering each of the successive frames of digitized analog speech information to synchronously generate time domain frame classification vectors at the fixed rate, wherein each of the time domain frame classification vectors is derived from one of the successive frames of digitized analog speech information, comprising

means for converting each of the successive frames of digitized analog speech information to a spectral domain vector using N multi-taper discrete prolate spheroid sequence basis (MTDPSSB) functions that are factors of a Fredholm integral of the first kind, wherein N is a positive integer; and

means for converting each spectral domain vector to one of the time domain frame classification vectors using Inverse Discrete Cosine Transformation, wherein N is; and

means for synchronously generating a sequence of a set of visemes wherein each set of visemes in the sequence is derived from a corresponding one of the time domain frame classification vectors, wherein each set of visemes is synchronously generated with a latency less than 10 milliseconds with reference to a successive frame of digitized analog speech information with which the set of visemes corresponds.

15. (Currently amended) The method for extracting visemes from an audio speech signal according to claim 1, wherein in the steps of conversion, N is 5 or less. ~~and wherein in the step of analyzing, and each set of visemes is generated with a latency less than 10 milliseconds with reference to a successive frame of digitized analog speech information with which the set of visemes corresponds~~.